



(19) Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) Publication number : **0 399 666 B1**

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification :
28.07.93 Bulletin 93/30

(51) Int. Cl.⁵ : **C12N 15/62, C07K 13/00,
C12P 21/02**

(21) Application number : **90304575.5**

(22) Date of filing : **26.04.90**

(54) **Fusion proteins containing N-terminal fragments of human serum albumin.**

Consolidated with 90907285.2/0470165
(European application No./publication No.) by
decision dated 20.07.92.

(30) Priority : **29.04.89 GB 8909919**

(43) Date of publication of application :
28.11.90 Bulletin 90/48

(45) Publication of the grant of the patent :
28.07.93 Bulletin 93/30

(84) Designated Contracting States :
AT BE CH DE DK ES FR GR IT LI LU NL SE

(56) References cited :
**EP-A- 0 308 381
EP-A- 0 322 094**

(73) Proprietor : **Delta Biotechnology Limited
Castle Court, Castle Boulevard
Nottingham NG7 1FD (GB)**

(72) Inventor : **Ballance, David James
11 South Road
West Bridgford, Nottingham NG2 7AG (GB)**

(74) Representative : **Bassett, Richard Simon et al
ERIC POTTER & CLARKSON St. Mary's Court
St. Mary's Gate
Nottingham NG1 1LE (GB)**

EP 0 399 666 B1

Note : Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

Description

The present invention relates to fusion polypeptides where two individual polypeptides or parts thereof are fused to form a single amino acid chain. Such fusion may arise from the expression of a single continuous coding sequence formed by recombinant DNA techniques.

Fusion polypeptides are known, for example those where a polypeptide which is the ultimately desired product of the process is expressed with an N-terminal "leader sequence" which encourages or allows secretion of the polypeptide from the cell. An example is disclosed in EP-A-116 201 (Chiron).

Human serum albumin (HSA) is a known protein found in the blood. EP-A-147 198 (Delta Biotechnology) discloses its expression in a transformed host, in this case yeast. Our earlier application EP-A-322 094 discloses N-terminal fragments of HSA, namely those consisting of residues 1-n where n is 369 to 419, which have therapeutic utility. The application also mentions the possibility of fusing the C-terminal residue of such molecules to other, unnamed, polypeptides.

One aspect of the present invention provides a fusion polypeptide comprising, as at least part of the N-terminal portion thereof, an N-terminal portion of HSA or a variant thereof and, as at least part of the C-terminal portion thereof, another polypeptide except that, when the said N-terminal portion of HSA is the 1-n portion where n is 369 to 419 or a variant thereof then the said polypeptide is (a) the 585 to 1578 portion of human fibronectin or a variant thereof, (b) the 1 to 368 portion of CD4 or a variant thereof, (c) platelet derived growth factor, or a variant thereof, (d) transforming growth factor, or a variant thereof, (e) the 1-261 portion of mature human plasma fibronectin or a variant thereof, (f) the 278-578 portion of mature human plasma fibronectin or a variant thereof, (g) the 1-272 portion of mature human von Willebrand's Factor or a variant thereof, or (h) alpha-1-antitrypsin or a variant thereof.

The N-terminal portion of HSA is preferably the said 1-n portion, the 1-177 portion (up to and including the cysteine), the 1-200 portion (up to but excluding the cysteine) or a portion intermediate 1-177 and 1-200.

The term "human serum albumin" (HSA) is intended to include (but not necessarily to be restricted to) known or yet-to-be-discovered polymorphic forms of HSA. For example, albumin Naskapi has Lys-372 in place of Glu-372 and pro-albumin Christchurch has an altered pro-sequence. The term "variants" is intended to include (but not necessarily to be restricted to) minor artificial variations in sequence (such as molecules lacking one or a few residues, having conservative substitutions or minor insertions of residues, or having minor variations of amino acid structure). Thus polypeptides which have 80%, preferably 85%, 90%, 95% or 99%, homology with HSA are deemed to be "variants". It is also preferred for such variants to be physiologically equivalent to HSA; that is to say, variants preferably share at least one pharmacological utility with HSA. Furthermore, any putative variant which is to be used pharmacologically should be non-immunogenic in the animal (especially human) being treated.

Conservative substitutions are those where one or more amino acids are substituted for others having similar properties such that one skilled in the art of polypeptide chemistry would expect at least the secondary structure, and preferably the tertiary structure, of the polypeptide to be substantially unchanged. For example, typical such substitutions include asparagine for glutamine, serine for asparagine and arginine for lysine. Variants may alternatively, or as well, lack up to ten (preferably only one or two) intermediate amino acid residues (ie not at the termini of the said N-terminal portion of HSA) in comparison with the corresponding portion of natural HSA; preferably any such omissions occur in the 100 to 369 portion of the molecule (relative to mature HSA itself) (if present). Similarly, up to ten, but preferably only one or two, amino acids may be added, again in the 100 to 369 portion for preference (if present). The term "physiologically functional equivalents" also encompasses larger molecules comprising the said sequence plus a further sequence at the N-terminal (for example, pro-HSA, pre-pro-HSA and met-HSA).

Clearly, the said "another polypeptide" in the fusion compounds of the invention cannot be the remaining portion of HSA, since otherwise the whole polypeptide would be HSA, which would not then be a "fusion polypeptide".

Even when the HSA-like portion is not the said 1-n portion of HSA, it is preferred for the non-HSA portion to be one of the said (a) to (h) entities.

The 1 to 368 portion of CD4 represents the first four disulphide-linked immunoglobulin-like domains of the human T lymphocyte CD4 protein, the gene for and amino acid sequence of which are disclosed in D. Smith *et al* (1987) *Science* 328, 1704-1707. It is used to combat HIV infections.

The sequence of human platelet-derived growth factor (PDGF) is described in Collins *et al* (1985) *Nature* 316, 748-750. Similarly, the sequence of transforming growth factors β (TGF- β) is described in Deryck *et al* (1985) *Nature* 316, 701-705. These growth factors are useful for wound-healing.

A cDNA sequence for the 1-261 portion of Fn was disclosed in EP-A-207 751 (obtained from plasmid pFH6 with endonuclease PvuII). This portion binds fibrin and can be used to direct fused compounds to blood clots.

A cDNA sequence for the 278-578 portion of Fn, which contains a collagen-binding domain, was disclosed by R.J. Owens and F.E. Baralle in 1986 E.M.B.O.J. 5, 2825-2830. This portion will bind to platelets.

The 1-272 portion of von Willebrand's Factor binds and stabilises factor VIII. The sequence is given in Bon-tham *et al*, Nucl. Acids Res. 14, 7125-7127.

5 Variants of alpha-1-antitrypsin include those disclosed by Rosenburg *et al* (1984) Nature 312, 77-80. In particular, the present invention includes the Pittsburgh variant (Met³⁵⁸ is mutated to Arg) and the variant where Pro³⁵⁷ and Met³⁵⁸ are mutated to alanine and arginine respectively. These compounds are useful in the treatment of septic shock and lung disorders.

10 Variants of the non-HSA portion of the polypeptides of the invention include variations as discussed above in relation to the HSA portion, including those with conservative amino acid substitutions, and also homologues from other species.

15 The fusion polypeptides of the invention may have N-terminal amino acids which extend beyond the portion corresponding to the N-terminal portion of HSA. For example, if the HSA-like portion corresponds to an N-terminal portion of mature HSA, then pre-, pro-, or pre-pro sequences may be added thereto, for example the yeast alpha-factor leader sequence. The fused leader portions of WO 90/01063 may be used. The polypeptide which is fused to the HSA portion may be a naturally-occurring polypeptide, a fragment thereof or a novel polypeptide, including a fusion polypeptide. For example, in Example 3 below, a fragment of fibronectin is fused to the HSA portion via a 4 amino acid linker.

20 It has been found that the amino terminal portion of the HSA molecule is so structured as to favour particularly efficient translocation and export of the fusion compounds of the invention in eukaryotic cells.

25 A second aspect of the invention provides a transformed host having a nucleotide sequence so arranged as to express a fusion polypeptide as described above. By "so arranged", we mean, for example, that the nucleotide sequence is in correct reading frame with an appropriate RNA polymerase binding site and translation start sequence and is under the control of a suitable promoter. The promoter may be homologous with or heterologous to the host. Downstream (3') regulatory sequences may be included if desired, as is known. The host is preferably yeast (for example Saccharomyces spp., e.g. S. cerevisiae; Kluyveromyces spp., e.g. K. lactis; Pichia spp.; or Schizosaccharomyces spp., e.g. S. pombe) but may be any other suitable host such as E. coli, B. subtilis, Aspergillus spp., mammalian cells, plant cells or insect cells.

30 A third aspect of the invention provides a process for preparing a fusion polypeptide according to the first aspect of the invention by cultivation of a transformed host according to the second aspect of the invention, followed by separation of the fusion polypeptide in a useful form.

35 A fourth aspect of the invention provides therapeutic methods of treatment of the human or other animal body comprising administration of such a fusion polypeptide.

In the methods of the invention we are particularly concerned to improve the efficiency of secretion of useful therapeutic human proteins from yeast and have conceived the idea of fusing to amino-terminal portions of HSA those proteins which may ordinarily be only inefficiently secreted. One such protein is a potentially valuable wound-healing polypeptide representing amino acids 585 to 1578 of human fibronectin (referred to herein as Fn 585-1578). As we have described in a separate application (filed simultaneously herewith) this molecule contains cell spreading, chemotactic and chemokinetic activities useful in healing wounds. The fusion polypeptides of the present invention wherein the C-terminal portion is Fn 585-1578 can be used for wound healing applications as biosynthesised, especially where the hybrid human protein will be topically applied. However, the portion representing amino acids 585 to 1578 of human fibronectin can if desired be recovered from the fusion protein by preceding the first amino acid of the fibronectin portion by amino acids comprising a factor X cleavage site. After isolation of the fusion protein from culture supernatant, the desired molecule is released by factor X cleavage and purified by suitable chromatography (e.g. ion-exchange chromatography). Other sites providing for enzymatic or chemical cleavage can be provided, either by appropriate juxtaposition of the N-terminal and C-terminal portions or by the insertion therebetween of an appropriate linker.

40 At least some of the fusion polypeptides of the invention, especially those including the said CD4 and vWF fragments, PDGF and α_1 AT, also have an increased half-life in the blood and therefore have advantages and therapeutic utilities themselves, namely the therapeutic utility of the non-HSA portion of the molecule. In the case of α_1 AT and others, the compound will normally be administered as a one-off dose or only a few doses over a short period, rather than over a long period, and therefore the compounds are less likely to cause an immune response.

55 EXAMPLES : SUMMARY

Standard recombinant DNA procedures were as described by Maniatis *et al* (1982 and recent 2nd edition) unless otherwise stated. Construction and analysis of phage M13 recombinant clones was as described by

Messing (1983) and Sanger et al (1977).

DNA sequences encoding portions of human serum albumin used in the construction of the following molecules are derived from the plasmids mHOB12 and pDBD2 (EP-A-322 094, Delta Biotechnology Ltd, relevant portions of which are reproduced below) or by synthesis of oligonucleotides equivalent to parts of this sequence. DNA sequences encoding portions of human fibronectin are derived from the plasmid pFHDEL1, or by synthesis of oligonucleotides equivalent to parts of this sequence. Plasmid pFHDEL1, which contains the complete human cDNA encoding plasma fibronectin, was obtained by ligation of DNA derived from plasmids pH6, 16, 54, 154 and 1 (EP-A-207 751; Delta Biotechnology Ltd).

This DNA represents an mRNA variant which does not contain the 'ED' sequence and had an 89-amino acid variant of the III-CS region (R.J. Owens, A.R. Kornblith and F.E. Baralle (1986) Oxford Surveys on Eukaryotic Genes 3 141-160). The map of this vector is disclosed in Fig. 11 and the protein sequence of the mature polypeptide produced by expression of this cDNA is shown in Fig. 5.

Oligonucleotides were synthesised on an Applied Biosystems 380B oligonucleotide synthesiser according to the manufacturer's recommendations (Applied Biosystems, Warrington, Cheshire, UK).

An expression vector was constructed in which DNA encoding the HSA secretion signal and mature HSA up to and including the 387th amino acid, leucine, fused in frame to DNA encoding a segment of human fibronectin representing amino acids 585 to 1578 inclusive, was placed downstream of the hybrid promoter of EP-A-258 067 (Delta Biotechnology), which is a highly efficient galactose-inducible promoter functional in Saccharomyces cerevisiae. The codon for the 1578th amino acid of human fibronectin was directly followed by a stop codon (TAA) and then the S. cerevisiae phosphoglycerate kinase (PGK) gene transcription terminator. This vector was then introduced into S. cerevisiae by transformation, wherein it directed the expression and secretion from the cells of a hybrid molecule representing the N-terminal 387 amino acids of HSA C-terminally fused to amino acids 585 to 1578 of human fibronectin.

In a second example a similar vector is constructed so as to enable secretion by S. cerevisiae of a hybrid molecule representing the N-terminal 195 amino acids of HSA C-terminally fused to amino acids 585 to 1578 of human fibronectin.

Aspects of the present invention will now be described by way of example and with reference to the accompanying drawings, in which:

Figure 1 (on two sheets) depicts the amino acid sequence currently thought to be the most representative of natural HSA, with (boxed) the alternative C-termini of HSA(1-n);
 Figure 2 (on two sheets) depicts the DNA sequence coding for mature HSA, wherein the sequence included in Linker 3 is underlined;
 Figure 3 illustrates, diagrammatically, the construction of mHOB16;
 Figure 4 illustrates, diagrammatically, the construction of pHOB31;
 Figure 5 (on 6 sheets) illustrates the mature protein sequence encoded by the Fn plasmid pFHDEL1;
 Figure 6 illustrates Linker 5, showing the eight constituent oligonucleotides;
 Figure 7 shows schematically the construction of plasmid pDBDF2;
 Figure 8 shows schematically the construction of plasmid pDBDF5;
 Figure 9 shows schematically the construction of plasmid pDBDF9;
 Figure 10 shows schematically the construction of plasmid DBDF12, using plasmid pFHDEL1; and
 Figure 11 shows a map of plasmid pFHDEL1.

EXAMPLE 1 : HSA 1-387 FUSED TO Fn 585-1578

The following is an account of a preparation of plasmids comprising sequences encoding a portion of HSA, as is disclosed in EP-A-322 094.

The human serum albumin coding sequence used in the construction of the following molecules is derived from the plasmid M13mp19.7 (EP-A-201 239, Delta Biotech- nology Ltd.) or by synthesis of oligonucleotides equivalent to parts of this sequence. Oligonucleotides were synthesised using phosphoramidite chemistry on an Applied Biosystems 380B oligonucleotide synthesizer according to the manufacturer's recommendations (AB Inc., Warrington, Cheshire, England).

An oligonucleotide was synthesised (Linker A) which represented a part of the known HSA coding sequence (Figure 2) from the PstI site (1235-1240, Figure 2) to the codon for valine 381 wherein that codon was changed from GTG to GTC:

Linker 1

		D	P	H	E	C	Y
5	5'	GAT	CCT	CAT	GAA	TGC	TAT
	3'	ACGT	CTA	GGA	GTA	CTT	ACG
1247							
10		A	K	V	F	D	E
15		GCC	AAA	GTG	TTC	GAT	GAA
		CGG	TTT	CAC	AAG	CTA	CTT
1267							
20		P	L	V			
		CTT	GTC	3'			
25		GGA	CAG	5'			

Linker 1 was ligated into the vector M13mp19 (Norrrander *et al.*, 1983) which had been digested with PstI and HincII and the ligation mixture was used to transfet E.coli strain XL1-Blue (Stratagene Cloning Systems, San Diego, CA). Recombinant clones were identified by their failure to evolve a blue colour on medium containing the chromogenic indicator X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) in the present of IPTG (isopropylthio-β-galactoside). DNA sequence analysis of template DNA prepared from bacteriophage particles of recombinant clones identified a molecule with the required DNA sequence, designated mHOB12 (Figure 3).

M13mp19.7 consists of the coding region of mature HSA in M13mp19 (Norrrander *et al.*, 1983) such that the codon for the first amino acid of HSA, GAT, overlaps a unique Xhol site thus:

		Asp Ala	
	5'	C T C G A G A T G C A	3'
40	3'	G A G C T C T A C G T	5'
<u>Xhol</u>			

45 (EP-A-210 239). M13mp19.7 was digested with Xhol and made flush-ended by S1-nuclease treatment and was then ligated with the following oligonucleotide (Linker 2):

Linker 2

50	5' T C T T T T A T C C A A G C T T G G A T A A A A G A 3'
	3' A G A A A A T A G G T T C G A A C C T A T T T C T 5'
<u>HindIII</u>	

The ligation mix was then used to transfet E.coli XL1-Blue and template DNA was prepared from several

plaques and then analysed by DNA sequencing to identify a clone, pDBD1 (Figure 4), with the correct sequence.

A 1.1 kb HindIII to PstI fragment representing the 5' end of the HSA coding region and one half of the inserted oligonucleotide linker was isolated from pDBD1 by agarose gel electrophoresis. This fragment was then ligated with double stranded mHOB12 previously digested with HindIII and PstI and the ligation mix was then used to transfect E.coli XL1-Blue. Single stranded template DNA was prepared from mature bacteriophage particles of several plaques. The DNA was made double stranded in vitro by extension from annealed sequencing primer with the Klenow fragment of DNA polymerase I in the presence of deoxynucleoside triphosphates. Restriction enzyme analysis of this DNA permitted the identification of a clone with the correct configuration, mHOB15 (Figure 4).

The following oligonucleotide (Linker 3) represents from the codon for the 382nd amino acid of mature HSA (glutamate, GAA) to the codon for lysine 389 which is followed by a stop codon (TAA) and a HindIII site and then a BamHI cohesive end:

15 Linker 3

	E	E	P	Q	N	L	I	K	J	
20										3'
	5' GAA GAG CCT CAG AAT TTA ATC AAA TAA GCTTG									
	3' CTT CTC GGA GTC TTA AAT TAG TTT ATT CGAACCTAG									5'

25 This was ligated into double stranded mHOB15, previously digested with HincII and BamHI. After ligation, the DNA was digested with HincII to destroy all non-recombinant molecules and then used to transfect E.coli XL1-Blue. Single stranded DNA was prepared from bacteriophage particles of a number of clones and subjected to DNA sequence analysis. One clone having the correct DNA sequence was designated mHOB16 (Figure 4).

30 A molecule in which the mature HSA coding region was fused to the HSA secretion signal was created by insertion of Linker 4 into BamHI and Xhol digested M13mp19.7 to form pDBD2 (Figure 4).

Linker 4

35

	M	K	W	V	S	F	
40	5' GATCC	ATG	AAG	TGG	GTA	AGC	TTT
	G	TAC	TTC	ACC	CAT	TCG	AAA
45	I	S	L	L	F	L	F
	ATT	TCC	CTT	CTT	TTT	CTC	TTT
	TAA	AGG	GAA	GAA	AAA	GAG	AGC
							TCG

50

55

	S	A	Y	S	R	G	V	F
5	TCG	GCT	TAT	TCC	AGG	GGT	GTG	TTT
	AGC	CGA	ATA	AGG	TCC	CCA	CAC	AAA
10	R	R						
	CG	3'						
	GCAGCT	5'						

15 In this linker the codon for the fourth amino acid after the initial methionine, ACC for threonine in the HSA pre-pro leader sequence (Lawn et al., 1981), has been changed to AGC for serine to create a HindIII site.
 A sequence of synthetic DNA representing a part of the known HSA coding sequence (Lawn et al., 1981) (amino acids 382 to 387, Fig. 2), fused to part of the known fibronectin coding sequence (Kornblhtt et al., 1985) (amino acids 585 to 640, Fig. 2), was prepared by synthesising six oligonucleotides (Linker 5, Fig. 6). The oligonucleotides 2, 3, 4, 6, 7 and 8 were phosphorylated using T4 polynucleotide kinase and then the oligonucleotides were annealed under standard conditions in pairs, i.e. 1+8, 2+7, 3+6 and 4+5. The annealed oligonucleotides were then mixed together and ligated with mHOB12 which had previously been digested with the restriction enzymes HincII and EcoRI. The ligation mixture was then used to transfet E.coli XL1-Blue (Stratagene Cloning Systems, San Diego, CA). Single stranded template DNA was then prepared from mature bacteriophage particles derived from several independent plaques and then was analysed by DNA sequencing. A clone in which a linker of the expected sequence had been correctly inserted into the vector was designated pDBDF1 (Fig. 7). This plasmid was then digested with PstI and EcoRI and the approx. 0.24kb fragment was purified and then ligated with the 1.29kb BamHI-PstI fragment of pDBD2 (Fig. 7) and BamHI + EcoRI digested pUC19 (Yanisch-Perron, et al., 1985) to form pDBDF2 (Fig. 7).

20 25 30 A plasmid containing a DNA sequence encoding full length human fibronectin, pFHDEL1, was digested with EcoRI and XbaI and a 0.77kb EcoRI-XbaI fragment (Fig. 8) was isolated and then ligated with EcoRI and Sall digested M13 mp18 (Norlander et al., 1983) to form pDBDF3 (Fig. 8).

35 The following oligonucleotide linker (Linker 6) was synthesised, representing from the PstI site at 4784-4791 of the fibronectin sequence of EP-A-207 751 to the codon for tyrosine 1578 (Fig. 5) which is followed by a stop codon (TAA), a HindIII site and then a BamHI cohesive end:

Linker 6

40	G	P	D	Q	T	E	M	T	I	E	G	L
	GGT	CCA	GAT	CAA	ACA	GAA	ATG	ACT	ATT	GAA	GGC	TTG
45	A	CGT	CCA	GGT	CTA	GTT	TGT	CTT	TAC	TGA	TAA	CTT
												CCG AAC
	Q	P	T	V	E	Y						Stop
50	CAG	CCC	ACA	GTG	GAG	TAT			TAA			GCTTG
	GTC	GGG	TGT	CAC	CTC	ATA			ATT			CGAACCTAG

55 This linker was then ligated with PstI and HindIII digested pDBDF3 to form pDBDF4 (Fig. 8). The following DNA fragments were then ligated together with BglII digested pKV50 (EP-A-258 067) as shown in Fig. 8: 0.68kb EcoRI-BamHI fragment of pDBDF4, 1.5kb BamHI-StuI fragment of pDBDF2 and the 2.2kb StuI-EcoRI fragment of pFHDEL1. The resultant plasmid pDBDF5 (Fig. 8) includes the promoter of EP-A-258 067 to direct the ex-

pression of the HSA secretion signal fused to DNA encoding amino acids 1-387 of mature HSA, in turn fused directly and in frame with DNA encoding amino acids 585-1578 of human fibronectin, after which translation would terminate at the stop codon TAA. This is then followed by the S.cerevisiae PGK gene transcription terminator. The plasmid also contains sequences which permit selection and maintenance in Escherichia coli and S.cerevisiae (EP-A-258 067).

This plasmid was introduced into S.cerevisiae S150-2B (leu2-3 leu2-112 ura3-52 trp1-289 his3-1) by standard procedures (Beggs, 1978). Transformants were subsequently analysed and found to produce the HSA-fibronectin fusion protein.

10 EXAMPLE 2 : HSA 1-195 FUSED TO Fn 585-1578

In this second example the first domain of human serum albumin (amino acids 1-195) is fused to amino acids 585-1578 of human fibronectin.

15 The plasmid pDBD2 was digested with BamHI and BglII and the 0.79kb fragment was purified and then ligated with BamHI-digested M13mp19 to form pDBDF6 (Fig. 6). The following oligonucleotide:

5'-C C A A A G C T C G A G G A A C T T C G-3'

20 was used as a mutagenic primer to create a Xhol site in pDBDF6 by in vitro mutagenesis using a kit supplied by Amersham International PLC. This site was created by changing base number 696 of HSA from a T to a G (Fig. 2). The plasmid thus formed was designated pDBDF7 (Fig. 9). The following linker was then synthesised to represent from this newly created Xhol site to the codon for lysine 195 of HSA (AAA) and then from the codon for isoleucine 585 of fibronectin to the ends of oligonucleotides 1 and 8 shown in Fig. 6.

25 Linker 7

D	E	L	R	D	E	G	K	A	S	S	A	K
TC GAT GAA CTT CGG GAT GAA GGG AAG GCT TCG TCT GCC AAA												
A CTT GAA GCC CTA CTT CCC TTC CGA AGC AGA CGG TTT												
I T E T P S Q P N S H												
ATC ACT GAG ACT CCG AGT CAG C												
TAG TGA CTC TGA GGC TCA GTC GGG TTG AGG GTG G												

35 This linker was ligated with the annealed oligonucleotides shown in Fig. 3, i.e. 2+7, 3+6 and 4+5 together with Xhol and EcoRI digested pDBDF7 to form pDBDF8 (Fig. 9). Note that in order to recreate the original HSA DNA sequence, and hence amino acid sequence, insertion of linker 7 and the other oligonucleotides into pDBDF7 does not recreate the Xhol site.

40 The 0.83kb BamHi-StuI fragment of pDBDF8 was purified and then was ligated with the 0.68kb EcoRI-BamHI fragment of pDBDF2 and the 2.22kb StuI-EcoRI fragment of pFHDEL1 into BglII-digested pKV50 to form pDBDF9 (Fig. 9). This plasmid is similar to pDBDF5 except that it specifies only residues 1-195 of HSA rather than 1-387 as in pDBDF5.

45 When introduced into S.cerevisiae S150-2B as above, the plasmid directed the expression and secretion of a hybrid molecule composed of residues 1-195 of HSA fused to residues 585-1578 of fibronectin.

EXAMPLE 3 : HSA 1-387 FUSED TO Fn 585-1578, AS CLEAVABLE MOLECULE

50 In order to facilitate production of large amounts of residues 585-1578 of fibronectin, a construct was made in which DNA encoding residues 1-387 of HSA was separated from DNA encoding residues 585-1578 of fibronectin by the sequence

I E G R
ATT GAA GGT AGA
TAA CTT CCA TCT

which specifies the cleavage recognition site for the blood clotting Factor X. Consequently the purified secreted product can be treated with Factor X and then the fibronectin part of the molecule can be separated from the HSA part.

To do this two oligonucleotides were synthesised and then annealed to form Linker 8.

Linker 8

15	E	E	P	Q	N	L	I	E	G
	GAA	GAG	CCT	CAG	AAT	TTA	ATT	GAA	GGT
20	CTT	CTC	GGA	GTC	TTA	AAT	TA	CTT	CCA
	R	I	T	E	T	P	S	Q	P
25	AGA	ATC	ACT	GAG	ACT	CCG	AGT	CAG	C
	TCT	TAG	TGA	CTC	TGA	GGC	TCA	GTC	GGG
30	N	S	H						
35	TTG	AGG	GTG	G					

This linker was then ligated with the annealed oligonucleotides shown in Fig. 6, i.e. 2+7, 3+6 and 4+5 into HincII and EcoRI digested mHOB12, to form pDBDF10 (Fig. 7). The plasmid was then digested with PstI and EcoRI and the roughly 0.24kb fragment was purified and then ligated with the 1.29kb BamHI-PstI fragment of pDBD2 and BamHI and EcoRI digested pUC19 to form pDBDF11 (Fig. 10).

The 1.5kb BamH1-StuI fragment of pDBDF11 was then ligated with the 0.68kb EcoRI-BamH1 fragment of pDBDF4 and the 2.22kb StuI-EcoRI fragment of pFHDEL1 into BglII-digested pKV50 to form pDBDF12 (Fig. 10). This plasmid was then introduced into S.cerevisiae S150-2B. The purified secreted fusion protein was treated with Factor X to liberate the fibronectin fragment representing residues 585-1578 of the native molecule.

REFERENCES

- Beggs, J.D. (1978) *Nature* 275, 104-109
50 Kornblith et al. (1985) *EMBO J.* 4, 1755-1759
Lawn, R.M. et al. (1981) *Nucl. Acid. Res.* 9, 6103-6114
Maniatis, T. et al. (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold
Spring Harbor, New York.
55 Messing, J. (1983) *Methods Enzymol.* 101, 20-78
Norrander, J. et al. (1983) *Gene* 26, 101-106
Sanger, F. et al. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467
Yanisch-Perron, C. (1985) *Gene* 33, 103-119

Claims**Claims for the following Contracting States : AT, BE, CH, LI, DE, DK, FR, IT, LU, NL, SE**

- 5 1. A fusion polypeptide comprising, as at least part of the N-terminal portion thereof, an N-terminal portion
of HSA or a variant thereof and, as at least part of the C-terminal portion thereof, another polypeptide
except that, when the said N-terminal portion of HSA is the 1-n portion where n is 369 to 419 or a variant
thereof then the said polypeptide is (a) the 585 to 1578 portion of human fibronectin or a variant thereof,
10 (b) the 1 to 368 portion of CD4 or a variant thereof, (c) platelet derived growth factor or a variant thereof,
(d) transforming growth factor β or a variant thereof, (e) the 1-261 portion of mature human plasma fi-
bronectin or a variant thereof, (f) the 278-578 portion of mature human plasma fibronectin or a variant
thereof, (g) the 1-272 portion of mature human von Willebrand's Factor or a variant thereof, or (h) alpha-
1-antitrypsin or a variant thereof.
- 15 2. A fusion polypeptide according to Claim 1 additionally comprising at least one N-terminal amino acid ex-
tending beyond the portion corresponding to the N-terminal portion of HSA.
- 20 3. A fusion polypeptide according to Claim 1 or 2 wherein there is a cleavable region at the junction of the
said N-terminal or C-terminal portions.
- 25 4. A fusion polypeptide according to any one of the preceding claims wherein the said C-terminal portion is
the 585 to 1578 portion of human plasma fibronectin or a variant thereof.
- 30 5. A transformed or transfected host having a nucleotide sequence so arranged as to express a fusion poly-
peptide according to any one of the preceding claims.
- 35 6. A process for preparing a fusion polypeptide by cultivation of a host according to Claim 5, followed by sep-
aration of the fusion polypeptide in a useful form.
- 40 7. A fusion polypeptide according to any one of Claims 1 to 4 for use in therapy.

Claims for the following Contracting States : ES, GR

- 45 1. A process for preparing a fusion polypeptide by (i) cultivation of a transformed or transfected host having
a nucleotide sequence so arranged as to express a fusion polypeptide, followed by (ii) separation of the
fusion polypeptide in a useful form, characterised in that the fusion polypeptide comprises as at least part
of the N-terminal portion thereof, an N-terminal portion of HSA or a variant thereof and, as at least part
of the C-terminal portion thereof, another polypeptide except that, when the said N-terminal portion of
HSA is the 1-n portion where n is 369 to 419 or a variant thereof then the said polypeptide is (a) the 585
to 1578 portion of human fibronectin or a variant thereof, (b) the 1 to 368 portion of CD4 or a variant there-
of, (c) platelet derived growth factor or a variant thereof, (d) transforming growth factor β or a variant there-
of, (e) the 1-261 portion of mature human plasma fibronectin or a variant thereof, (f) the 278-578 portion
of mature human plasma fibronectin or a variant thereof, (g) the 1-272 portion of mature human von Will-
ebrand's Factor or a variant thereof, or (h) alpha-1-antitrypsin or a variant thereof.
- 50 2. A process according to Claim 1, wherein the fusion polypeptide additionally comprising at least one N-
terminal amino acid extending beyond the portion corresponding to the N-terminal portion of HSA.
- 55 3. A process according to Claim 1 or 2 wherein, in the fusion polypeptide, there is a cleavable region at the
junction of the said N-terminal or C-terminal portions.
- 60 4. A process according to any one of the preceding claims wherein the said C-terminal portion is the 585
to 1578 portion of human plasma fibronectin or a variant thereof.

Patentansprüche**Patentansprüche für folgende Vertragsstaaten : AT, BE, CH, DE, DK, FR, IT, LU, NL, SE**

- 5 1. Fusionspolypeptid, umfassend als mindestens einen Teil seines N-terminalen Teils einen N-terminalen Teil von HSA oder eine Variante davon und als mindestens einen Teil seines C-terminalen Teils ein weiteres Polypeptid mit der Ausnahme, daß wenn es sich bei dem N-terminalen Teil von HSA um den Teil 1-n mit n = 369 bis 419 oder eine Variante davon handelt, das Polypeptid aus
- 10 (a) dem Teil 585 bis 1578 von Humanfibronectin oder einer Variante davon,
 (b) dem Teil 1 bis 368 von CD4 oder einer Variante davon,
 (c) dem "Platelet Derived Growth Factor" (PDGF) oder einer Variante davon,
 (d) dem "Transforming Growth Factor β " (TGF β) oder einer Variante davon,
 (e) dem Teil 1-261 von reifem Humanplasmafibronectin oder einer Variante davon,
 (f) dem Teil 278-578 von reifem Humanplasmafibronectin oder einer Variante davon,
 (g) dem Teil 1-272 von reifem Human-von Willebrand's-Faktor oder einer Variante davon oder
 (h) Alpha-1-Antitrypsin oder einer Variante davon, besteht.
- 15 2. Fusionspolypeptid nach Anspruch 1, zusätzlich umfassend mindestens eine N-terminale Aminosäure, die länger als der dem N-terminalen Teil von HSA entsprechende Teil ist.
- 20 3. Fusionspolypeptid nach Anspruch 1 oder 2, bei dem sich an der Verbindung der N-terminalen oder C-terminalen Teile eine spaltbare Region befindet.
- 25 4. Fusionspolypeptid nach einem der vorhergehenden Ansprüche, wobei der C-terminale Teil aus dem Teil 585 bis 1578 von Humanplasmafibronectin oder einer Variante davon besteht.
- 30 5. Transformierter oder transfizierter Wirt mit einer Nukleotidsequenz, die so angeordnet ist, daß sie ein Fusionspolypeptid nach einem der vorhergehenden Ansprüche exprimieren kann.
- 35 6. Verfahren zur Herstellung eines Fusionspolypeptids durch Kultivieren eines Wirts nach Anspruch 5 und anschließendes Abtrennen des Fusionspolypeptids in einer geeigneten Form.
- 40 7. Fusionspolypeptid nach einem der Ansprüche 1 bis 4 zur therapeutischen Verwendung.

Patentansprüche für folgende Vertragsstaaten : ES, GR

- 35 1. Verfahren zur Herstellung eines Fusionspolypeptids durch
- 40 (i) Kultivieren eines transformierten oder transfektierten Wirts mit einer Nukleotidsequenz, die so angeordnet ist, daß sie ein Fusionspolypeptid exprimiert, und
 dadurch gekennzeichnet, daß das Fusionspolypeptid als mindestens einen Teil seines N-terminalen Teils einen N-terminalen Teil von HSA oder eine Variante davon und als mindestens einen Teil seines C-terminalen Teils ein weiteres Polypeptid umfaßt, mit der Ausnahme, daß wenn es sich bei dem N-terminalen Teil von HSA um den Teil 1-n mit n= 369 bis 419 oder eine Variante davon handelt, das Polypeptid aus
- 45 (a) dem Teil 585-1578 von Humanfibronectin oder einer Variante davon,
 (b) dem Teil 1-368 von CD4 oder einer Variante davon,
 (c) dem Platelet Derived Growth Factor oder einer Variante davon,
 (d) dem Transforming Growth Factor β oder einer Variante davon,
 (e) dem Teil 1-261 von reifem Humanplasmafibronectin oder einer Variante davon,
 (f) dem Teil 278-578 von reifem Humanplasmafibronectin oder einer Variante davon,
 (g) dem Teil 1-272 von reifem Human-von Willebrand's-Faktor oder einer Variante davon oder
 (h) α -1-Antitrypsin oder einer Variante davon besteht.
- 50 2. Verfahren nach Anspruch 1, wobei das Fusionspolypeptid zusätzlich mindestens eine N-terminale Aminosäure, die länger als der dem N-terminalen Teil von HSA entsprechende Teil ist, umfaßt.
- 55 3. Verfahren nach Anspruch 1 oder 2, wobei sich in dem Fusionspolypeptid an der Verbindung der N-terminalen oder C-terminalen Teile eine spaltbare Region befindet.

4. Verfahren nach einem der vorhergehenden Ansprüche, wobei der C-terminale Teil aus dem Teil 585-1578 von Humanplasmafibronectin oder einer Variante davon besteht.

5 Revendications

Revendications pour les Etats contractants suivants : AT, BE, CH, DE, DK, FR, IT, LU, NL, SE

1. Polypeptide fusionné comprenant en tant qu'au moins une partie de sa portion N-terminale, une portion N-terminale de HSA ou d'un variant de celle-ci et, en tant qu'au moins une partie de sa portion C-terminale, un autre polypeptide sauf que, lorsque cette portion N-terminale de HSA est la portion 1-n dans laquelle n est 369 à 419 ou un variant de celle-ci, ce polypeptide est (a) la portion 585 à 1578 de la fibronectine humaine ou un variant de celle-ci, (b) la portion 1 à 368 de CD4 ou un variant de celle-ci, (c) le facteur de croissance dérivé des plaquettes sanguines ou un variant de celui-ci, (d) le facteur de croissance β de transformation ou un variant de celui-ci, (e) la portion 1-261 de la fibronectine mature de plasma humain ou un variant de celle-ci, (f) la portion 278-578 de la fibronectine mature de plasma humain ou un variant de celle-ci, (g) la portion 1-272 du facteur humain mature de von Willebrand ou un variant de celle-ci, ou (h) l'alpha-1-antitrypsine ou un variant de celle-ci.
2. Polypeptide fusionné suivant la revendication 1, comprenant de plus au moins un acide aminé N-terminal se prolongeant au-delà de la portion correspondant à la portion N-terminale de HSA.
3. Polypeptide fusionné suivant les revendications 1 ou 2, dans lequel il y a une région susceptible d'être coupée à la jonction de ces portions N-terminale et C-terminale.
4. Polypeptide fusionné suivant l'une quelconque des revendications précédentes, dans lequel cette portion C-terminale est la portion 585 à 1578 de la fibronectine de plasma humain ou un variant de celle-ci.
5. Hôte transformé ou transfecté ayant une séquence de nucléotides arrangée de façon à exprimer un polypeptide fusionné suivant l'une quelconque des revendications précédentes.
6. Procédé pour préparer un polypeptide fusionné par culture d'un hôte suivant la revendication 5, suivie de la séparation du polypeptide fusionné sous une forme utile.
7. Polypeptide fusionné suivant l'une quelconque des revendications 1 à 4 utilisable en thérapie.

Revendications pour les Etats contractants suivants : ES, GR

1. Procédé pour préparer un polypeptide fusionné par (i) la culture d'un hôte transformé ou transfecté ayant une séquence de nucléotides arrangée de façon à exprimer un polypeptide fusionné, suivie de (ii) la séparation du polypeptide fusionné sous une forme utile, caractérisé en ce que le polypeptide fusionné comprend, en tant qu'au moins une partie de sa portion N-terminale, une portion N-terminale de HSA ou d'un variant de celle-ci et, en tant qu'au moins une partie de sa portion C-terminale, un autre polypeptide sauf que, lorsque cette portion N-terminale de HSA est la portion 1-n dans laquelle n est 369 à 419 ou un variant de celle-ci, ce polypeptide est alors (a) la portion 585 à 1578 de la fibronectine humaine ou un variant de celle-ci, (b) la portion 1 à 368 de CD4 ou un variant de celle-ci, (c) le facteur de croissance dérivé des plaquettes sanguines ou un variant de celui-ci, (d) le facteur de croissance β de transformation ou un variant de celui-ci, (e) la portion 1-261 de la fibronectine mature de plasma humain ou un variant de celle-ci, (f) la portion 278-578 de la fibronectine mature de plasma humain ou un variant de celle-ci, (g) la portion 1-272 du facteur humain mature de von Willebrand ou un variant de celle-ci, ou (h) l'alpha-1-antitrypsine ou un variant de celle-ci.
2. Procédé suivant la revendication 1, dans lequel le polypeptide fusionné comprend de plus au moins un acide aminé N-terminal se prolongeant au-delà de la portion correspondant à la portion N-terminale de HSA.
3. Procédé suivant les revendications 1 ou 2 dans lequel, dans le polypeptide fusionné, il y a une région susceptible d'être coupée à la jonction de ces portions N-terminale et C-terminale.

4. Procédé suivant l'une quelconque des revendications précédentes, dans lequel cette portion C-terminale est la portion 585 à 1578 de la fibronectine de plasma humain ou un variant de celle-ci.

5

10

15

20

25

30

35

40

45

50

55

FIGURE 1

FIGURE 1 Cont.

Vai	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu	Gln	Leu	Gly	Glu
390																			
Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Vai	Arg	Tyr	Thr	Lys	Val	Pro	Gln	Val	Ser	Thr	
410																			
Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys	Val	Gly	Ser	Lys	Cys	Cys	Lys	His
430																			
Pro	Glc	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu	Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu
450																			
Cys	Val	Leu	His	Glu	Lys	Thr	Pro	Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys	Thr	Glu	Ser
470																			
Leu	Val	Asn	Arg	Arg	Pro	Cys	Phe	Ser	Ala	Leu	Glu	Val	Asp	Glu	Thr	Tyr	Val	Pro	Lys
490																			
Glu	Phe	Asn	Ala	Glu	Thr	Phe	Thr	Phe	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	
510																			
Arg	Gln	Ile	Lys	Lys	Gln	Thr	Ala	Leu	Val	Glu	Leu	Vai	Lys	His	Lys	Pro	Lys	Ala	
530																			
Lys	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	Cys	Cys	
550																			
Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	Lys	Leu	Val	Ala	Ala	Ser	
570																			
Ala	Ala	Leu	Gly	Leu															
590																			

FIGURE 2 DNA sequence coding for mature HSA

10	20	30	40	50	60	70	80
GATGCACACAAGACTGAGGTTGCTCATCGGTTAAAGATTGGGAGAAGAAAATTCAAAAGCCTGGTGTGATTGCCTT	D A H K S E V A H R F K D L G E E N F K A L V L I A F						
90	100	110	120	130	140	150	160
TGCTCAGTATCTTCAGCAGTGTCCATTGAAAGATCATGTAAGAAATTAGTGAATGAAAGTAAGTGAATTTGCCAAAAACATGTG	A Q Y L Q Q C P F E D H V K L V N E V T E F A K T C						
170	180	190	200	210	220	230	240
TTGCTGATGAGTCAGCTGAAATTGTGACPAATCACTTCATACCCCTTTTGGAGACAAATTATGCACAGTTGCAACTCTT	V A D E S A E N C D K S L H T I F G D K L C T V A T L						
250	260	270	280	290	300	310	320
CGTGAAACCTATGGTGAATGGCTGACTGCTGTGCAAAACAAAGAACCTGAGAGAAATGAAATGCTTCTTGCACACACPAAGA	R E T Y G E M A D C C A K Q E P E R N E C F L Q H K D						
330	340	350	360	370	380	390	400
TGACAAACCCAAACCTCCCCCATTGGTGAGACCAAGAGCTGATGTGATGTGCACTGCTTTCATGACAAATGAAGAGACAT	D N P N L P R L V R P E V D V M C T A F H D N E E T						
410	420	430	440	450	460	470	480
TTTGAAAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTATGCCCGGAACTCTTTCTTGCTAAAGG	F L K X Y L Y E I A R R H P Y F Y A P E L L F F A K R						
490	500	510	520	530	540	550	560
TATAAAGCTGCTTTACAGAAATGTTGCCAAGCTGCTGATAAAGCTGCCCTGCTGTTGCCAAAGCTCGATGAACTTCGGGA	Y K A A F T E C C Q A A D K A A C L L P K L D E L R D						
570	580	590	600	610	620	630	640
TGAAGGGAAAGGCTTCGCTGCCAACAGAGACTCAAATGTGCCAGTCTCCAAAAATTGGAGAAAAGAGCTTCAAGC	E G K A S S A K Q R L K C A S L Q K F G E R A F K A						
650	660	670	680	690	700	710	720
GGGCAGTGGCTCGCCTGAGCCAGAGATTCCCAGCTGAGTTGCAAGAGTTCCAAGTTAGTAGTGACAGATCTTACCAA	W A V A R L S Q R F P K A E F A E V S K L V T D L T K						
730	740	750	760	770	780	790	800
GTCCACACGGAAATGCTGCCATGGAGATCTGCTTGAATGTGCTGATGACAGGGCGGACCTGCCAACTATATCTGTGAAAAA	V H T E C C H G D L E C A D D R A D L A K Y I C E N						
810	820	830	840	850	860	870	880
TCAGGATTGATCTCCAGTAAACTGAAGGAATGCTGTGAAAAACCTCTGTTGGAAAAATCCCACGTGCATTGCCGAAGTGG	Q D S I S S K L K E C C E K P L L E K S H C I A E V						
890	900	910	920	930	940	950	960
AAAAATGATGAGATGCCCTGCTGACTTGCCTTCATTAGCTGCTGATTTGTTGAAAGTAAGGATGTTGCAAAACTATGCT	E N D E M P A D L P S L A A D F V E S K D V C K N Y A						
970	980	990	1000	1010	1020	1030	1040
GAGGCCAAAGGATGTCCTGGCATGTTTGTATGAAATATGCAAGAAGGCATCCTGATTACTCTGTGCTGCTGCTGCT	E A K D V F L G M F L Y E Y A R R H P D Y S V V L L L						

FIGURE 2 Cont.

1050 1060 1070 1080 1090 1100 1110 1120
 GAGACTTGCCAAGACATATGAAACCACCTAGAGAAGTGCCTGCCGCTGCAGATCCTCATGAATGCTATGCCAAAGTGT
 R L A K T Y E T T L E K C C A A A D P H E C Y A K V

 1130 1140 1150 1160 1170 1180 1190 1200
 TCGATGAATTAAACCTCTTGTGGAAAGGCCCTAGAATTAACTCAAACAAACTCTGAGCTTTTGAGCAGCTGGAGAG
 F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E

 1210 1220 1230 1240 1250 1260 1270 1280
 TACAAATTCCAGAATGCCTATTAGTCCTAACCAAGAAAGTACCCCCAAGTGTCAACTCCAACCTTGTAGAGGTCTC
 Y K F Q N A L L V R Y T K K V P Q V S T P T L V E V S

 1290 1300 1310 1320 1330 1340 1350 1360
 AAGAAACCTAGGAAAAGTGGGCACCAATGTTGTAACATCCTGAAGCAAAAGAAATGCCCTGTGCAGAAGACTATCTAT
 R N L G X V G S K C C K H P E A K R M P C A E D Y L

 1370 1380 1390 1400 1410 1420 1430 1440
 CCGTGGTCTGAAACCAGTTATGTTGCATGAGAAAACGCCAGTAAGTGCACAGAGTCACAAAATGCTCACAGACTCC
 S V V L N Q L C V L H E K T P V S D R V T K C C T E S

 1450 1460 1470 1480 1490 1500 1510 1520
 TTGGTGAACAGGGCACXTGCTTTCAAGCTCTGGAAAGTCGATGAAACATACGTCCAAAGAGTTAAATGCTGAACATT
 L V N R R P C F S A L E V D E T Y V P K E F N A E T F

 1530 1540 1550 1560 1570 1580 1590 1600
 CACCTTCCATGCAGATATAATGCACACTTCTGAGAAGGGAGAGACAAATCAAGAAACAAACTGCCACTTGTGAGCTTG
 T F H A D I C T L S E K E R Q I K K Q T A L V E L V

 1610 1620 1630 1640 1650 1660 1670 1680
 AACACAAAGCCCAGGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGATTCGCAGCTTGTAGAGAAGTGCAG
 K H X P K A T K E Q L K A V M D D F A A F V E K C C K

 1690 1700 1710 1720 1730 1740 1750 1760
 GCTGACGATAAGGAGACCTGCTTGGCCAGGGAGGGTAAAAAAACTTGTGCTGCAAGTCAGCTGCCCTAGGCTTATAACA
 A D D K E T C F A E E G K K L V A A S Q A A L G L

 1770 1780
 TCTACATTAAAAGCATCTCAG

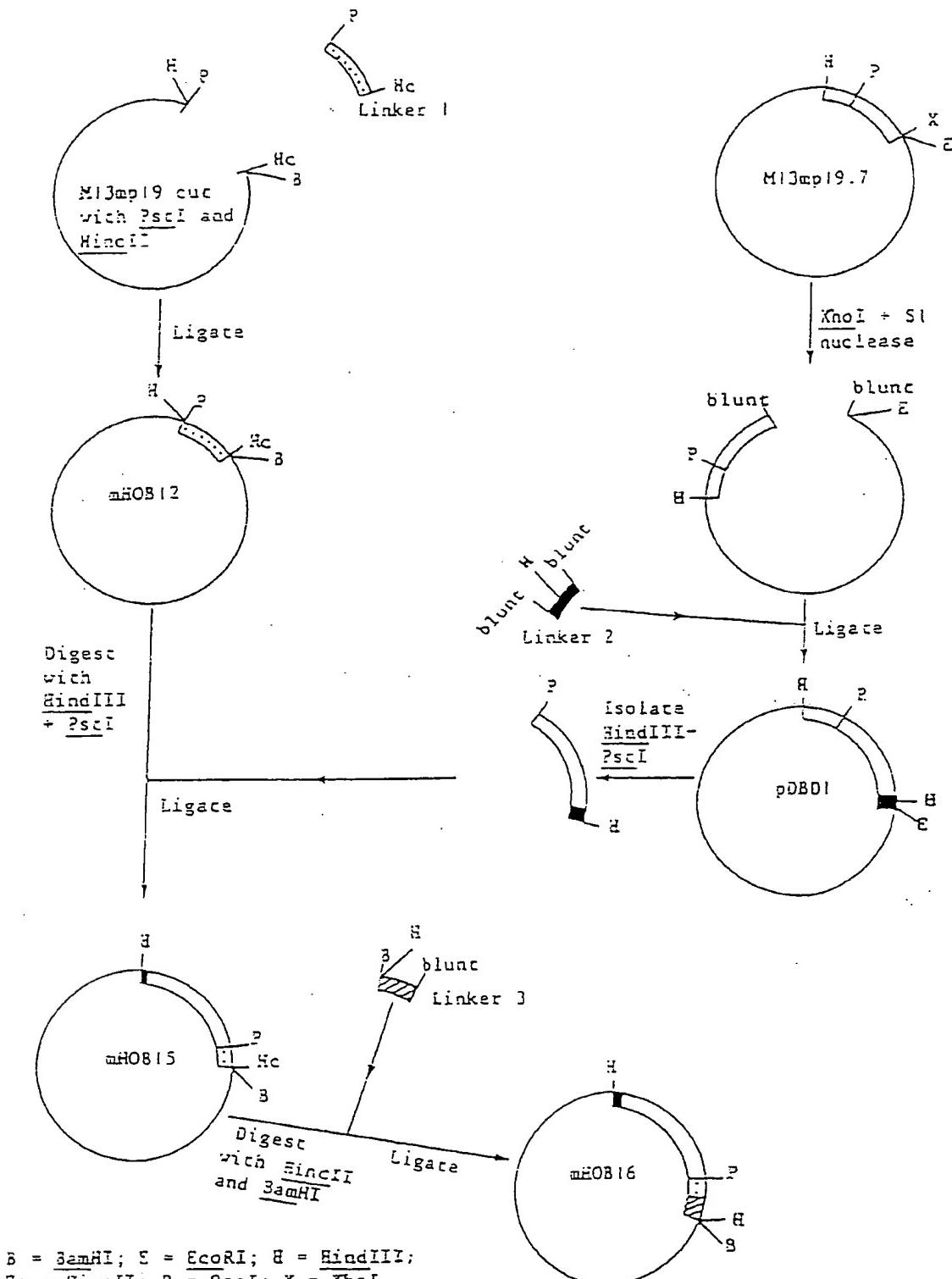
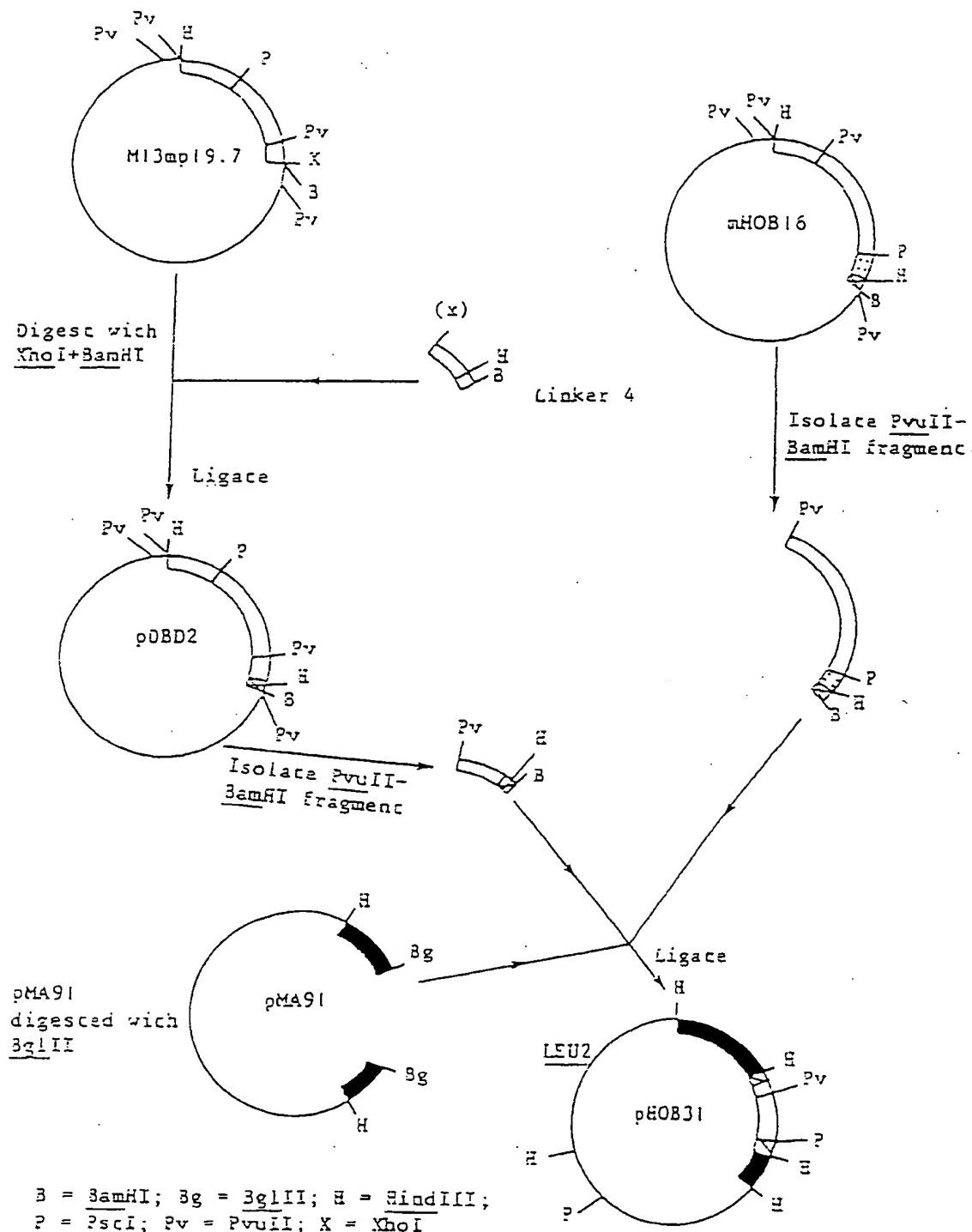
FIGURE 3 Construction of mHOB16

FIGURE 4 Construction of pE0831

Gin Ala Gin Gin Met Val Gin Pro Gin Ser¹⁰ Pro Val Ala Val Ser Gin Ser Lys Pro Gly²⁰
 Cys Tyr Asn Gly Lys His Tyr Gin³⁰ Ile Asn Gin Gin Trp Glu Arg Thr Tyr Leu Gly⁴⁰
 Asn Val Leu Val Cys Thr Cys Tyr Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro⁵⁰
 Glu Ala Glu Glu Thr Cys Phe Asp Lys Tyr⁶⁰ Thr Gly Asn Thr Tyr Arg Val Gly Asp Thr⁷⁰
 Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys Thr Cys Ile Gly Ala Gly Arg Gly⁸⁰
 Arg Ile Ser Cys Thr Ile Ala Asn Arg Cys His Glu Gly Gly Ser Tyr Lys Ile Gly⁹⁰
 Asp Thr Trp Arg Arg Pro His Glu Thr Gly¹⁰⁰ Gly Tyr Met Leu Glu Cys Val Cys Leu Gly¹¹⁰
 Asn Gly Lys Glu Trp Thr Cys Lys Pro Ile¹²⁰ Ala Glu Lys Cys Phe Asp His Ala Ala
 Gly Thr Ser Tyr Val Val Gly Glu Thr Trp Glu Lys Pro Tyr Gin Gly Trp Met Met Val¹³⁰
 Asp Cys Thr Cys Leu Gly Glu Ser Gly Arg Ile Thr Cys Thr Ser Arg Asn Arg Cys¹⁴⁰
 Asn Asp Gln Asp Thr Arg Thr Ser Tyr Arg Ile¹⁵⁰ Gly Asp Thr Trp Ser Lys Asp Asn¹⁶⁰
 Arg Gly Asn Leu Leu Gin Cys Ile Cys Thr Gly Asn Gly Arg Gly Glu Trp Lys Cys Glu¹⁷⁰
 Arg His Thr Ser Val Gin Thr Thr Ser Ser Gly Ser Gly Pro Pro Thr Asp Val Arg¹⁸⁰
 Ala Val Tyr Gin Pro His Pro Gin Pro Pro Tyr Gly His Cys Val Thr Asp¹⁹⁰
 Ser Ely Val Val Tyr Ser Val Gly Met Gin Trp Leu Lys Thr Gin Gly Asn Lys Gin²⁰⁰
 Leu Cys Thr Cys Leu Gly Asn Gly Val Ser Cys Gin Glu Thr Ala Val Thr Thr Tyr²¹⁰
 Gly Gly Asn Ser Asn Gly Glu Pro Cys Val²²⁰ Leu Pro Phe Thr Asn Gly Arg Thr Phe²³⁰
 Tyr Ser Cys Thr Thr Glu Gly Arg Gin Asp Gly His Leu Trp Cys Ser Thr Thr Ser Asn²⁴⁰
 Tyr Glu Glu Asp Gln Lys Tyr Ser Phe Cys Thr Asp His Thr Val Leu Val Gin Thr²⁵⁰ Gin²⁶⁰
 Gly Gly Asn Ser Asn Ely Ala Leu Cys His Phe Pro Phe Leu Tyr Asn Asn His Asn Tyr²⁷⁰
 Thr Asp Cys Thr Ser Glu Gly Arg Arg Asp Asn Met Lys Trp Cys Gly Thr Thr Gin Asn²⁸⁰
³⁹⁰ Asp Asn Met Lys Trp Cys Gly Thr Thr Gin Asn⁴⁰⁰
⁴¹⁰ Asp Asn Met Lys Trp Cys Gly Thr Thr Gin Asn⁴²⁰

Fig. 5A

Tyr Asp Ala Asp Gln Lys Phe Gly Phe Cys Pro Met Ala Ala His Glu Glu Ile Cys Thr
 Thr Asn Glu Gly Val Met Tyr Arg Ile Gly Asp Gln Trp Asp Lys Gln His Asp Met 460
 His Met Met Arg Cys Thr Cys Val Gly Asn Gly Arg Gly Glu Trp Thr Cys Tyr Ala Tyr
 Ser Gln Leu Arg Asp Gln Cys Ile Val Asp Asp Ile Thr Tyr Asn Val Asp Thr 480
 His Lys Arg His Glu Glu Gly His Met Leu Asn Cys Thr Cys Phe Gly Gln Gly Arg Phe
 Arg Trp Lys Cys Asp Pro Val Asp Gln Gln Asp Ser Glu Thr Gly Thr Phe Tyr 500
 Ile Gly Asp Ser Trp Glu Lys Tyr Val His Gly Val Arg Tyr Gln Cys Tyr Cys Tyr 520
 Arg Gly Ile Gly Glu Trp His Cys Gln Pro Leu Gln Thr Tyr Pro Ser Ser Gly 540
 Val Elu Val Phe Ile Thr Glu Thr Pro Ser Gln Pro Asn Ser His Pro Ile Gln Trp Asn
 Ala Pro Gln Pro Ser His Ile Ser Lys Ile Leu Arg Trp Arg Pro Lys Asn Ser Val 560
 Gly Arg Trp Lys Glu Ala Thr Ile Pro Gln His Leu Asn Ser Tyr Thr Ile Lys Gly Leu
 Lys Pro Gly Val Val Tyr Glu Gly Gln Ile Ser Ile Gln Gln Tyr Gly His Gln Glu 580
 Val Thr Arg Phe Asp Phe Thr Thr Thr Ser Thr Ser Thr Pro Val Thr Ser Asn Thr Val 600
 Thr Gly Glu Thr Thr Pro Phe Ser Pro Leu Val Ala Thr Ser Glu Ser Val Thr Glu Ile
 Thr Ala Ser Ser Phe Val Val Ser Trp Val Ser Ala Ser Asp Thr Val Ser Gly Phe Arg 620
 Val Glu Tyr Glu Leu Ser Glu Glu Asp Glu Pro Gln Tyr Leu Asp Leu Pro Ser Thr 640
 Ala Thr Ser Val Asn Ile Pro Asp Leu Leu Pro Ely Arg Lys Tyr Ile Val Asn Val Tyr
 Gln Ile Ser Glu Asp Gly Glu Gln Ser Leu Ile Leu Ser Thr Ser Gln Thr Thr Ala Pro 660
 Asp Ala Pro Asp Pro Thr Val Asp Gln Val Asp Asp Thr Ser Ile Val Val Arg Trp 680
 Ser Arg Pro Gln Ala Pro Ile Thr Gly Tyr Arg Ile Val Tyr Ser Pro Ser Val Glu Gly 700
 Ser Ser Thr Glu Leu Leu Pro Glu Thr Ala Asn Ser Val Thr Leu Ser Asp Leu Glu 820
 FNDEL 1 840

Fig. 5B

Pro Gly Val Gin Tyr Asn Ile Thr Ile Tyr Ala Val Glu Glu Asn Gln Glu Ser Thr Pro 850
 Val Val Ile Gin Gin Glu Thr Thr Gly Thr Pro Arg Ser Asp Thr Val Pro Ser Pro Pro 860
 Asp Leu Gin Phe Val Glu Val Thr Asp Val Lys Val Thr Ile Met Trp Thr Pro Pro 870
 Ser Ala Val Thr Gly Tyr Arg Val Asp Val Ile Pro Val Asn Leu Pro Gly Glu His Glu 880
 Gin Arg Leu Pro Ile Ser Arg Asn Thr 930 Phe Ala Glu Val Thr Gly Leu Ser Pro Gly Val 890
 Thr Tyr Tyr Phe Lys Val Phe Ala Val Ser His Gly Arg Glu Ser Lys Pro Leu Thr Ala 900
 Gin Gin Thr Thr Lys Leu Asp Ala Pro Thr Asn Leu Gin Phe Val Asn Glu Thr Asp Ser 910
 Thr Val Leu Val Arg Trp Thr Pro Pro Arg Ala Glu Ile Thr Gly Tyr Arg Leu Thr Val 920
 Gly Leu Thr Arg Arg Gly Gin Pro Arg Gin Tyr Asn Val Gly Pro Ser Val Ser Lys Tyr 930
 Pro Leu Arg Asn Leu Gin Pro Ala Ser 950 Gin Tyr Thr Val Ser Leu Val Ala Ile Lys Gly 940
 Asn Gin Glu Ser Pro Lys Ala Thr Gly Val Phe Thr Thr Leu Gin Pro Gly Ser Ser Ile 950
 Pro Pro Tyr Asn Thr Glu Val Thr Glu Thr Thr Ile Val Ile Thr Trp Thr Pro Ala Pro 960
 Arg Ile Gly Phe Lys Leu Gly Val Arg Pro Ser Gln Gly Gly Ala Pro Arg Glu Val 970
 Thr Ser Asp Ser Gly Ser Ile Val Val Ser Gly Leu Thr Pro Gly Val Glu Tyr Val Tyr 980
 Thr Ile Gin Val Leu Arg Asp Gly Gln Glu Arg Asp Ala Pro Ile Val Ash Lys Val Val 990
 Thr Pro Leu Ser Pro Pro Thr Asn Leu His Leu Glu Ala Asn Pro Asp Thr Gly Val Leu 1000
 Thr Val Ser Trp Glu Arg Ser Thr Thr Gly Arg Ile Thr Thr Thr Val 1010
 Pro Thr Asn Gly Gin Gln Gly Asn Ser Leu Glu Val Val His Ala Asp Gln Ser Ser 1020
 Cys Thr Phe Asp Asn Leu Ser Pro Gly Leu Glu Tyr Asn Val Ser Val Tyr Thr Val Val 1030
 ASP ASP Lys Glu Ser Val Pro Ile Ile Ser ASP Thr Ile Pro Ala Val Pro Pro Pro Thr 1040
 Asp Leu Arg Phe Thr Asn Ile Gly Pro Asp Thr Met Arg Val Thr Trp Ala Pro Pro Pro 1050

Fig. 5C

Ser Ile Asp Leu Thr Asn Phe Leu Val 1270 Arg Tyr Ser Pro Val Lys Asn Glu Glu Asp Val
 Ala Glu Leu Ser Ile Ser Pro Ser Asp Asn Ala Val Val Leu Thr Asn Leu Leu Pro Gly 1300
 Thr Glu Tyr Val Val Ser Val Ser Val Tyr Glu Gin His Glu Ser Thr Pro Leu Arg 1320
 Gly Arg Gin Lys Thr Gly Leu Asp Ser Pro Thr Gly Ile Asp Phe Ser Asp Ile Thr Ala 1340
 Asn Ser Phe Thr Val His Trp Ile Ala 1350 Pro Arg Ala Thr Ile Thr Gly Tyr Arg Ile Arg 1360
 His Pro Glu His Phe Ser Gly Arg Pro Arg Glu Asp Arg Val Pro His Ser Arg Asn 1380
 Ser Ile Thr Leu Thr Asn Lau Lau Thr Pro Gly Thr Glu Tyr Val Val Ser Ile Val Ala Leu 1400
 Asn Gly Arg Glu Glu Ser Pro Leu Leu Ile Gly Gin Gin Ser Thr Val Ser Asp Val Pro 1420
 Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr Ser Leu Leu Ile Ser Trp Asp Ala Pro 1440
 Ala Val Thr Val Arg Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val 1460
 Gin Glu Phe Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro Gly 1480
 Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp Ser Pro Ala Ser Ser 1500
 Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gin Met Gin Val Thr 1520
 Asp Val Gin Asp Asn Ser Ile Ser Val 1530 Lys Trp Leu Pro Ser Ser Pro Val Thr Gly 1540
 Tyr Arg Val Thr Thr Pro Lys Asn Gly Pro Gly Pro Thr Lys Thr Lys Thr Ala Gly 1560
 Pro Asp Gin Thr Glu Met Thr Ile Glu Glu Leu Gin Pro Thr Val Glu Tyr Val Val Ser 1580
 Val Tyr Ala Gin Asn Pro Ser Gly Glu Ser Gin Pro Leu Val Gin Thr Ala Val Thr Thr 1600
 Ile Pro Ala Pro Thr Asp Leu Lys Phe Thr Gin Val Thr Pro Thr Ser Leu Ser Ala Gin 1620
 Trp Thr Pro Pro Asn Val Gin Leu Thr Gly Tyr Arg Val Arg Val Thr Pro Lys Glu Lys 1640
 Thr Gly Pro Met Lys Glu Ile Asn Leu Ala Pro Asp Ser Ser Val Val Ser Gly 1660
 Leu Met Val Ala Thr Lys Tyr Glu Val Ser Val Tyr Ala Leu Lys Asp Thr Leu Thr Ser 1680
 FNDL 1

Fig. 5D

Arg Pro Ala Gin Gly Val Val Thr Thr Leu Glu Asn Val Ser Pro Pro Arg Arg Ala Arg
 Val Thr Asp Ala Thr Glu Thr Thr Ile Thr Ile Ser Trp Arg Thr Lys Thr Glu Thr Ile 1690
 Thr Gly Phe Gin Val Asp Ala Val Pro Ala Asn Gly Gin Thr Pro Ile Gin Arg Thr Ile 1710
 Lys Pro Asp Val Arg Ser Tyr Thr Ile Thr Gly Leu Gin Pro Gly Thr Asp Tyr Lys Ile 1730
 Tyr Leu Tyr Thr Leu Asn Asp Asn Ala Arg Ser Ser Pro Val Val Ile Asp Ala Ser Thr Ile 1750
 Ala Ile Asp Ala Pro Ser Asn Leu Arg Phe Leu Ala Thr Thr Pro Asn Ser Leu Leu Val 1770
 Ser Trp Gin Pro Pro Arg Ala Arg Ile Thr Gly Tyr Ile Ile Lys Tyr Glu Lys Pro Gly 1790
 Ser Pro Pro Arg Glu Val Val Pro Arg Pro Gly Val Thr Glu Ala Thr Ile Thr Ile 1810
 Gly Leu Glu Pro Gly Thr Glu Tyr Thr Ile Tyr Val Ile Ala Leu Lys Asn Asn Gin Lys 1830
 Ser Glu Pro Leu Ile Gly Arg Lys Lys Thr Asn Glu Leu Pro Gin Leu Val Thr Leu Pro 1850
 His Pro Asn Leu His Gly Pro Glu Ile Leu Asp Val Pro Ser Thr Val Gin Lys Thr Pro 1870
 Phe Val Thr His Pro Gly Tyr Asp Thr Gly Asn Gly Ile Gin Leu Pro Gly Thr Ser Gly 1890
 Gin Gin Pro Ser Val Gly Gin Met Ile Phe Glu Glu His Gly Phe Arg Arg Thr Thr 1910
 Pro Pro Thr Thr Ala Thr Pro Ile Arg His Arg Pro Tyr Pro Pro Asn Val Ala 1930
 Leu Ser Gin Thr Thr Ile Ser Trp Ala Pro Phe Gin Asn Thr Ser Glu Tyr Ile Ile Ser 1950
 Cys His Pro Val Gly Thr Asp Glu Glu Pro Leu Gin Phe Arg Val Pro Gly Thr Ser Thr 1970
 Ser Ala Thr Leu Thr Gly Leu Thr Arg Gly Ala Thr Tyr Asn Ile Val Glu Ala Leu 1990
 Lys Asp Gin Gin Arg His Lys Val Arg Elu Glu Val Val Thr Val Gly Asn Ser Val Asn 2010
 Glu Gly Leu Asn Gin Pro Thr Asp Asp Ser Cys Phe Asp Pro Tyr Thr Val Ser His Tyr 2030
 Ala Val Gly Asp Glu Trp Glu Arg Met Ser Gin Ser Gly Phe Lys Leu Leu Cys Gin Cys 2050
 Leu Ser Phe Gly Ser Gly His Phe Arg Cys Asp Ser Ser Arg Trp Cys His Asp Asn Gly 2070
 2090 2100

Fig. 5E

Val Asn Tyr Lys Ile Gly Glu Lys Trp Asp Arg Gln Gly Glu Asn Gly Gln Met Met Ser²¹¹⁰
 Cys Thr Cys Leu Gly Asn Gly Lys Gly Glu²¹³⁰ Phe Lys Cys Asp Pro His Glu Ala Thr Cys²¹⁴⁰
 Tyr Asp Asp Gly Lys Thr Tyr His Val Gly Glu Gln Trp Gln Lys Glu Tyr Leu Gly Ala²¹⁵⁰
 Ile Cys Ser Cys Thr Cys Phe Gly Gly²¹⁷⁰ Arg Gly Trp Arg Cys ASP Asn Cys Arg Arg²¹⁶⁰
 Pro Gly Glu Pro Ser Pro Glu Gly²¹⁹⁰ Thr Thr Gly Gln Ser Tyr Asn Gln Tyr Ser Gln²²⁰⁰
 Arg Tyr His Gln Arg Thr Asn Thr Asn Val Asn Cys Pro Ile Glu Cys Phe Met Pro Leu²²¹⁰
 Asp Val Gln Ala Asp Arg Glu Asp Ser Arg Glu²²³⁰

Fig. 5F

GAAGAGCCTCAGAATTAACTGAGACTCCGAGTCAGCCCACCTCCCACCCATCCAGTGG
CTTCTCGGAGTCTTAAATTAGTGAECTGAGGCTCAGTCGGTTGAGGGTGGGTAGGTCAACC

e e p q n l i t e | t p s q p n s h p i q w
8

AATGCACCACAGCCATCTCACATTCCAAGTACATTCTCAGGTGGAGACCTAAAAATTCTGTA
TTACGTGGTGTGGTAGAGTGTAAAGGTTCATGTAAGAGTCCACCTCTGGATTTTAAGACAT

n a p q p | s h i s k y i l r w r p k n s v
7

GGCCGTTGGAAGGAAGCTACCATACCAAGGCCACTTAAACTCCTACACCATCAAAGGCCTG
CCGGCAACCTTCCTCGATGGTATGGTCCGGTGAATTGAGGATGTGGTAGTTCCGGACTTAA

g | r w k e a t i p g h l n s | y t i k g l
6

Figure 6 Linker 5 showing the eight constituent oligonucleotides

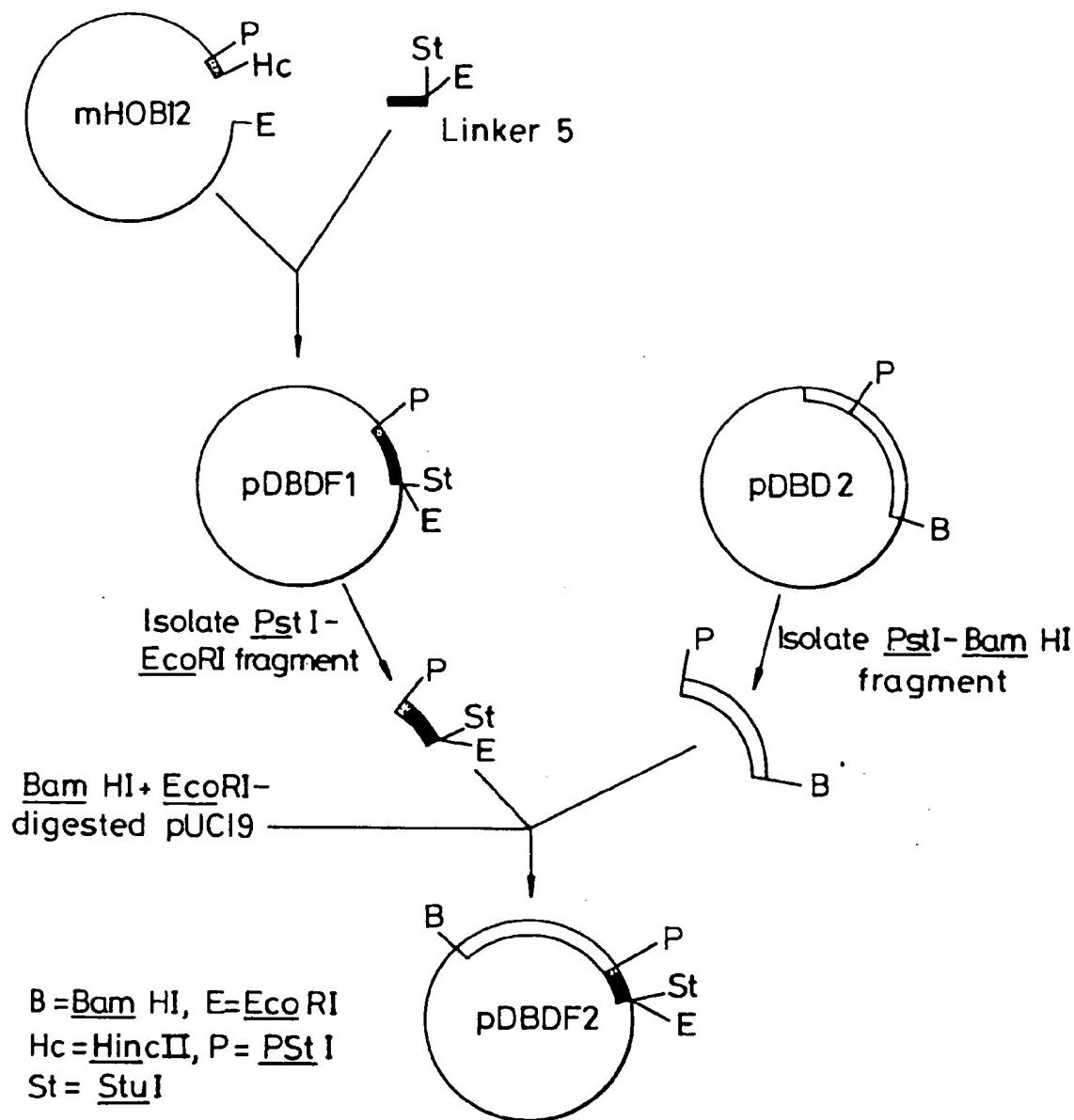


Fig. 7 Construction of pDBDF2

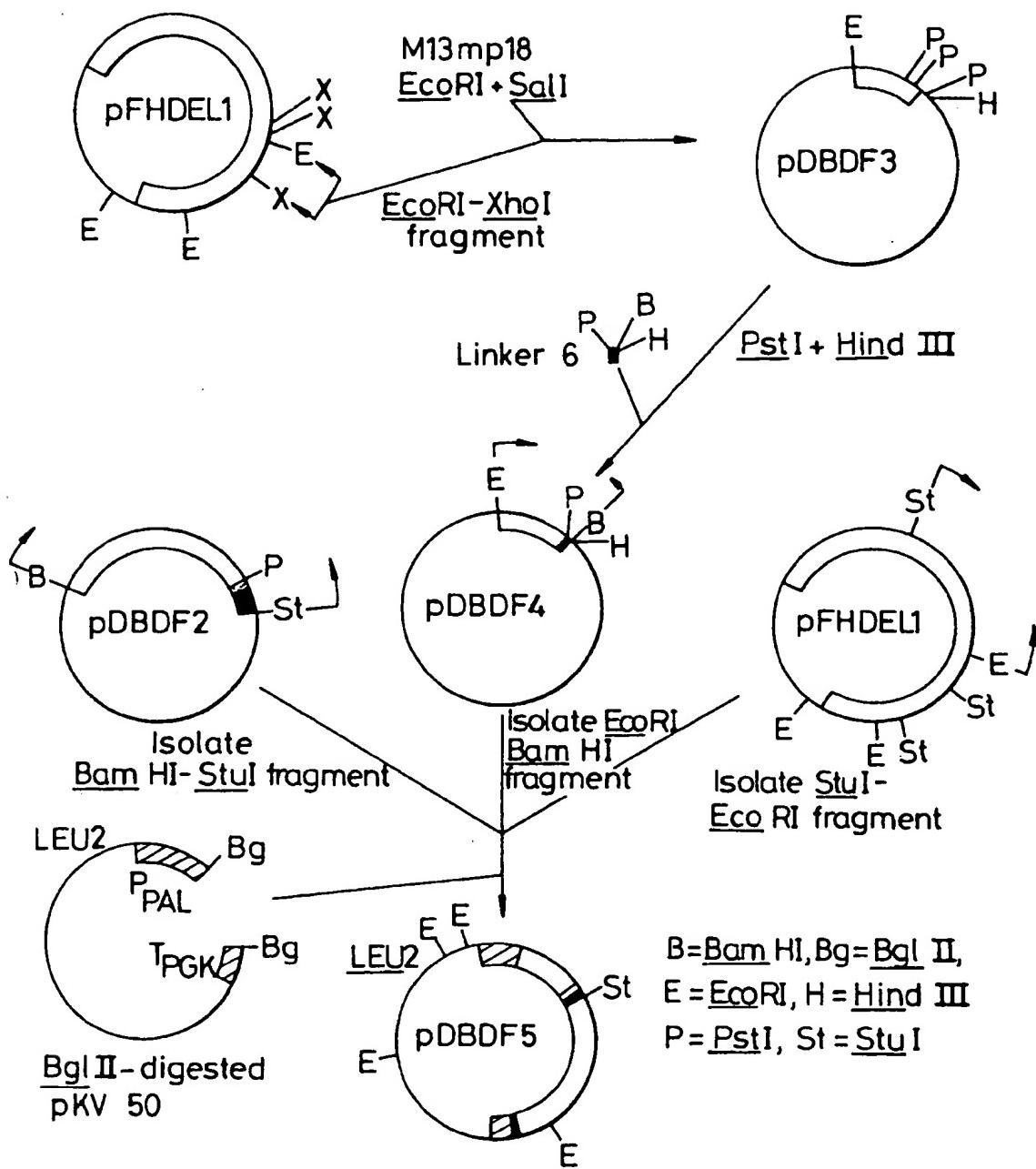


Fig. 8 Construction of pDBDF5

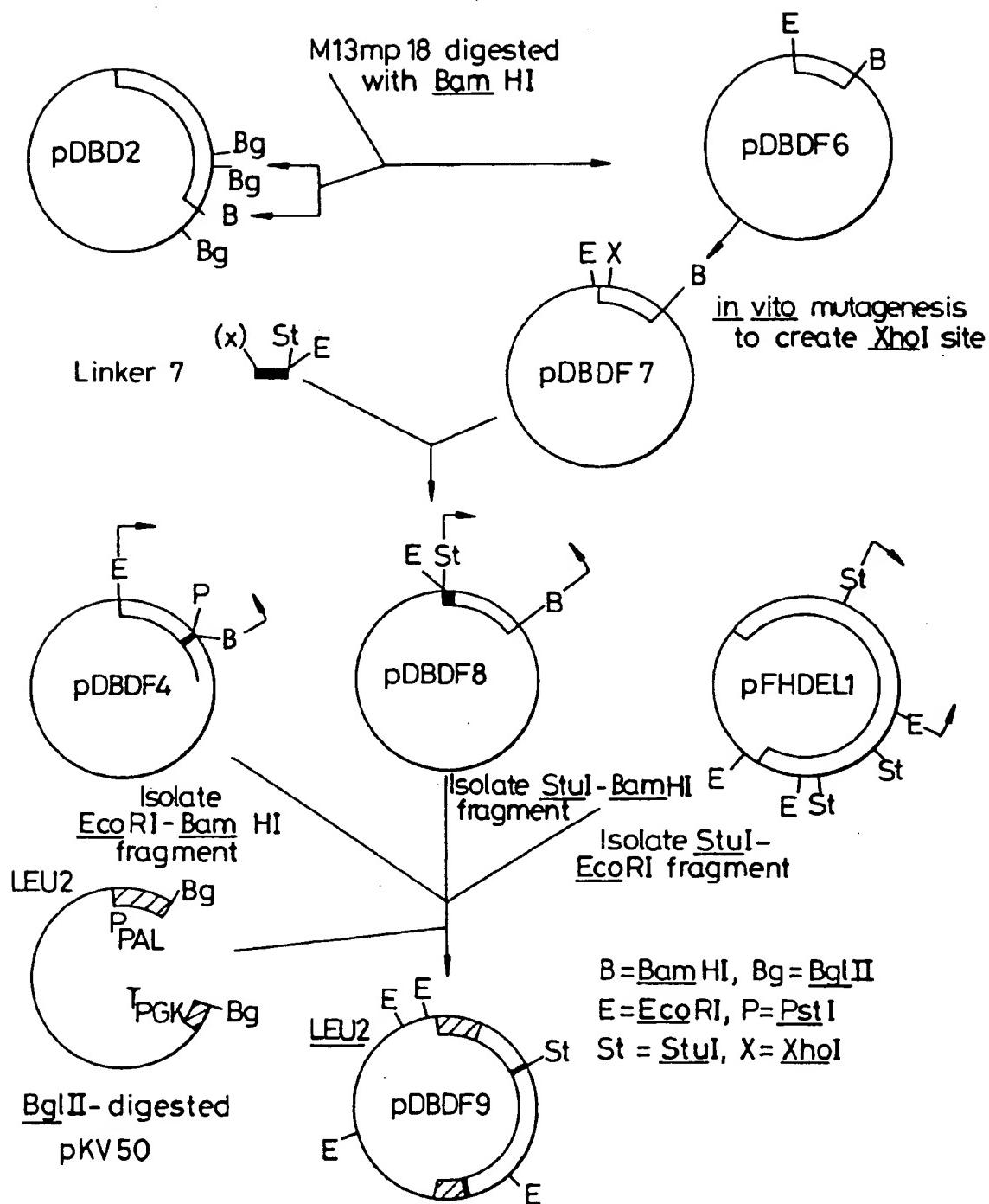


Fig. 9 Construction of pDBDF9

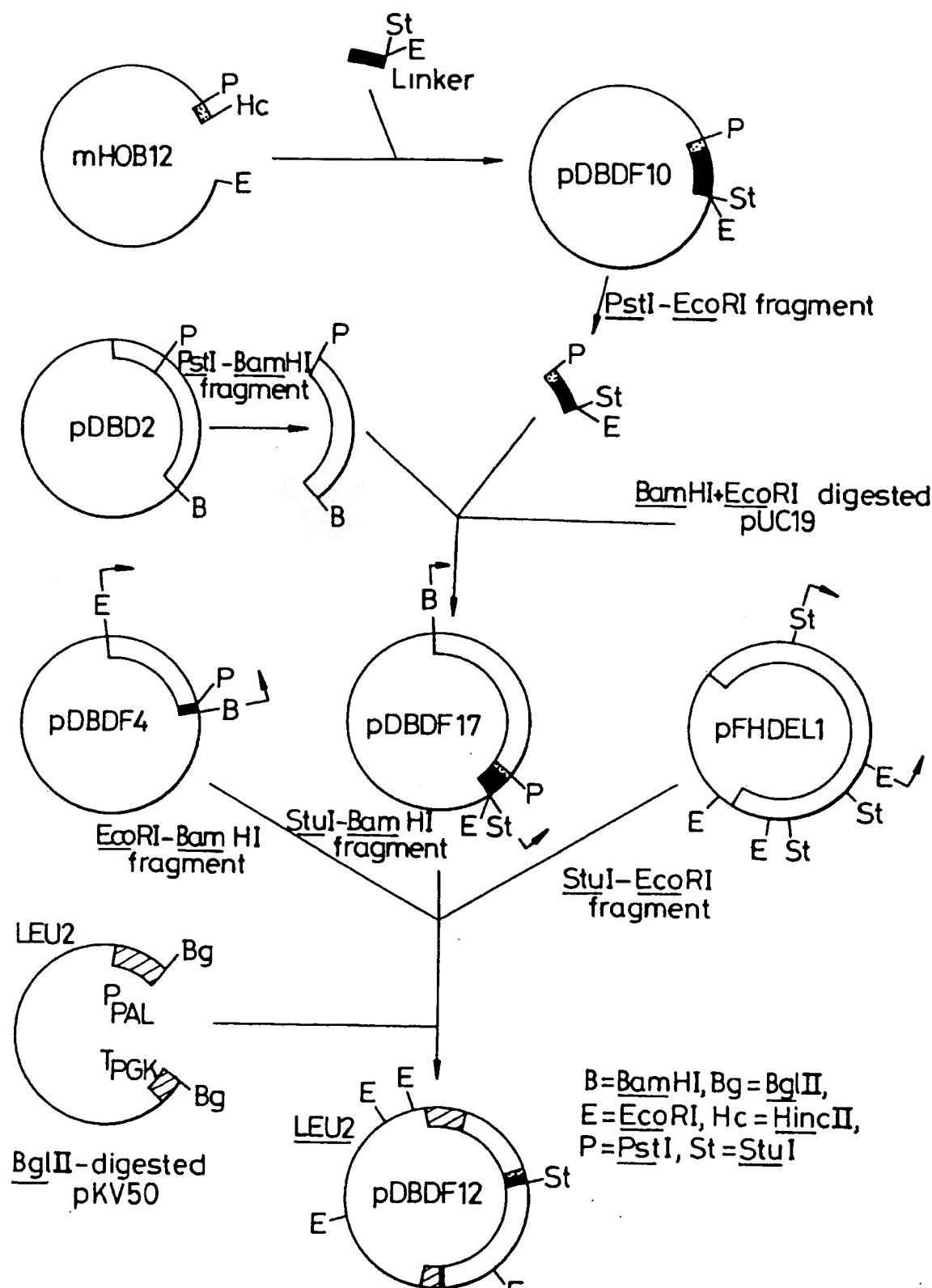


Fig. 10 Construction of pDBDF12

Figure 11

Name: pFHDEL1
Vector: pUC18 Amp^r 2860bp
Insert: hFNcDNA ~ 7630bp

